CLAIMS

What is claimed is:

1. A method for enhancing voice activity detection comprising:
determining a peak-to-mean likelihood ratio; and
comparing the peak-to-mean likelihood ratio to a selected threshold to determine whether a current audio frame represents a voice signal.

2. The method of claim 1, wherein prior to determining the peak-to-mean likelihood ratio, the method further comprises:
determining a short-term averaged energy for the current audio frame; and
determining a long-term averaged energy for the current audio frame.

3. The method of claim 2, wherein after determining the short-term averaged energy and the long-term averaged energy, the method further comprises:
determining whether a sum of the short-term averaged energy and a factor is greater than the long-term averaged energy; and
determining that the current audio frame represents silence if the sum is less than the long-term averaged energy, without necessitating a determination of the peak-to-mean likelihood ratio.

4. The method of claim 3, upon determining that the sum is greater than the long-term averaged energy and before determining the peak-to-mean likelihood ratio, the method further comprises:

determining whether a difference between the long-term averaged energy and the short-term averaged energy is less than a predetermined threshold;

determining that the current audio frame represents voice if the difference is greater than the predetermined threshold; and

continuing by determining the peak-to-mean likelihood ratio if the difference is less than the predetermined threshold.

5. The method of claim 2, wherein the determining of the short-term averaged energy comprises:

determining an energy, in decibels, of the current audio frame;

determining a short-term averaged energy for a prior audio frame; and

conducting a weighted average of the energy of the current audio frame and the short-term averaged energy for the prior audio frame.

6. The method of claim 1, wherein the determining a peak-to-mean likelihood ratio comprises

calculating an averaged peak-to-mean ratio for the current audio frame;

determining a maximum averaged peak-to-mean ratio;

determining a minimum averaged peak-to-mean ratio;

determining a first result being a difference between the maximum averaged peak-to-mean ratio and the averaged peak-to-mean ratio for the current audio frame;

determining a second result being a difference between the maximum averaged peak-to-mean ratio and the minimum averaged peak-to-mean ratio; and

conducting a ratio between the first result and the second result to produce the peak-to-mean likelihood ratio.

7. A communication module comprising:

a substrate;

a processing unit placed on the substrate; and

a memory coupled to the processing unit, the memory to contain a voice activity detector which, when executed by the processing unit, analyzes a short-term averaged energy, a long-term averaged energy, and a peak-to-mean likelihood ratio in order to determine whether a current audio frame represents voice or silence.

8. The communication module of claim 7, wherein the voice activity detector, when executed, controls the processing unit to determine whether a sum of the short-term averaged energy and a predetermined factor is greater than the long-term averaged energy, and to signal that the current audio frame represents silence if the sum is less than the long-term averaged energy.

9. The communication module of claim 8, wherein the voice activity detector, when executed, controls the processing unit to determine

whether a difference between the long-term averaged energy and the short-term averaged energy is less than a predetermined threshold, and to signal that the current audio frame represents voice if the difference is greater than the predetermined threshold.

10. The communication module of claim 9, wherein the voice activity detector, when executed, controls the processing unit to determine the peak-to-mean likelihood ratio, and to compare the peak-to-mean likelihood ratio to a selected threshold to determine whether a current audio frame represents a voice signal.

11. The communication module of claim 10, wherein the voice activity detector, when executed, controls the processing unit to determine a peak-to-mean ratio by (i) sampling an analog signal a predetermined number of times to produce a plurality of sampled signals each having a sampled value, (ii) determining a maximum value of the plurality of sampled signals, and (iii) conducting a ratio between an absolute value of the maximum value and a summation of the sampled values for the plurality of sampled signals.

12. The communication module of claim 10, wherein the voice activity detector, when executed, controls the processing unit to determine an averaged peak-to-mean ratio for the current audio frame by (i) monitoring a maximum averaged peak-to-mean ratio and a minimum averaged peak-to-mean ratio, (ii) determining a first result being a difference between the maximum averaged peak-to-mean ratio and the averaged peak-to-mean ratio for the current audio frame, (iii) determining a second result being a difference between the maximum averaged peak-to-mean ratio and the

minimum averaged peak-to-mean ratio, and (iv) conducting a ratio between the first result and the second result to produce the peak-to-mean likelihood ratio.

13. A machine readable medium having embodied thereon a computer program for processing by a machine, the computer program comprising:

a first routine for determining a peak-to-mean likelihood ratio; and

a second routine for comparing the peak-to-mean likelihood ratio to a selected threshold to determine whether an audio frame being transmitted represents a voice signal.

14. The machine readable medium of claim 13, wherein the computer program further comprising:

a third routine for determining a short-term averaged energy for the audio frame, the third routine being executed before the first and second routines; and

a fourth routine for determining a long-term averaged energy for the audio frame, the fourth routine being executed before the first and second routines.

15. The machine readable medium of claim 14, wherein the computer program further comprising:

a fifth routine for determining whether a sum of the short-term averaged energy and a predetermined factor is greater than the long-term averaged energy, the fifth routine being executed before the first and second routines; and

a sixth routine for determining whether a difference between the long-term averaged energy and the short-term averaged energy is less than a predetermined threshold, the sixth routine being executed after determining that the sum is greater than the long-term averaged energy and before execution of the first and second routines.

16. The machine readable medium of claim 15, wherein the fifth routine determining that the current audio frame represents silence if the sum is less than the long-term averaged energy.

17. The machine readable medium of claim 15, wherein the sixth routine determining that the current audio frame represents voice if the difference is greater than the predetermined threshold.

18. A voice activity detector comprising:

circuitry to determine a short-term averaged energy for an audio frame;

circuitry to determine a long-term averaged energy for the audio frame;

circuitry to determine whether the short-term averaged energy is greater than the long-term averaged energy by a predetermined factor;

circuitry to determine whether a difference between the long-term averaged energy and the short-term averaged energy is less than a predetermined threshold when the short-term averaged energy is greater than the long-term averaged energy by the predetermined factor;

circuitry to determine a peak-to-mean likelihood ratio when the difference between the long-term averaged energy and the short-term averaged energy is less than the predetermined threshold; and

circuitry to comparing the peak-to-mean likelihood ratio to a selected threshold and to determine that the audio frame represents a voice signal when the peak-to-mean likelihood ratio is greater than a selected threshold.

add a1

add C8

Add E7